Assessment of ChatGPT Generated Patient Education Materials in Head and Neck Surgery

Lana Mnajjed, MS3; Rusha Patel MD, FACS
OU College of Medicine – Tulsa, OU Department of Otolaryngology - Head and Neck Surgery

INTRODUCTION

- Artificial intelligence (AI) is a rapidly expanding field that medicine is incorporating into practice. AI has already been used successfully for clinical diagnosis, surgery, drug development, medical management, and medical education.
- ChatGPT is an AI chatbot that functions in a much more advanced manner than the previous technology similar to it. It can hold a continuous conversation, share accurate information, and give life-like responses when prompted.
- ChatGPT is free to use and widely available to the public and could potentially be used to provide counseling to patients in between and in addition to provider appointments. This can fill in gaps of patient knowledge due to poor information retention and unaddressed concerns/issues during visits. However, the adequacy of this type of counseling remains unknown.
- Our aim was to assess ChatGPT generated patient education with publicly available material with regards to 5 standard clinical scenarios encountered in head and neck oncologic surgery.

METHODS

- ChatGPT access was obtained through institutional email registration.
- 5 Common head and neck procedures were used for assessment: laryngectomy, parotidectomy, thyroidectomy, neck dissection, and glossectomy.
- Publicly available patient education material was searched for using Google (search: procedure name followed by "patient information. Where available, the most comprehensive information from four academic institutions was selected for comparison.
- ChatGPT was presented with a set series of questions regarding the above clinical scenarios (ex: " "I'm going to have a(n) X surgery. Can you tell me more about it?". Responses were recorded and saved.
- All material was analyzed using 2 methods: the Patient Education Materials Assessment Tool (PEMAT) and the Suitability Assessment of Materials (SAM). Only relevant categories were used for scoring. A readability consensus of 7 readability formulas was used to assess readability of the chat and materials when assessed by the SAM, and the SAM and PEMAT guidelines were used for all other respective scoring values.

Figure 1: Laryngectomy Material Scores

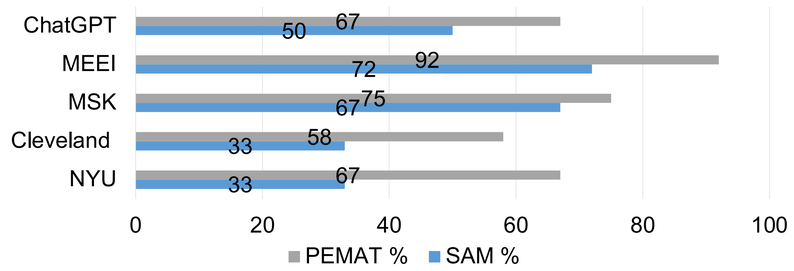


Figure 2: Parotidectomy Material Scores

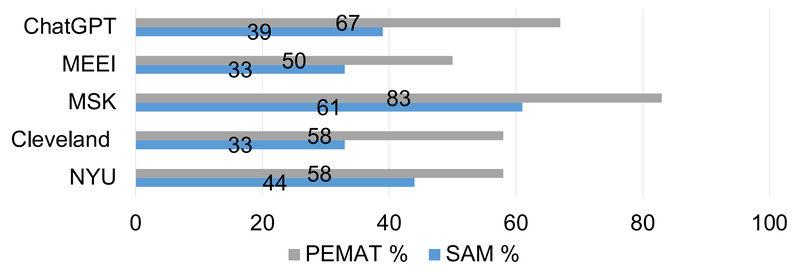


Figure 3: Thyroidectomy Material Scores

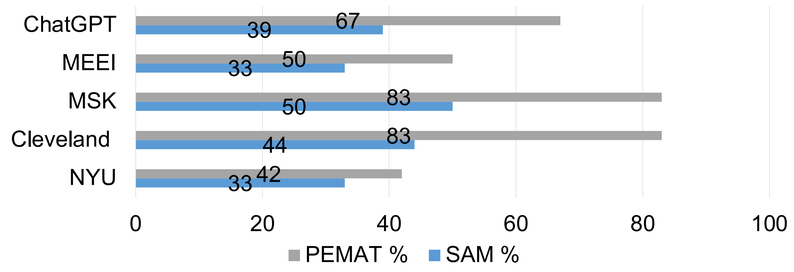


Figure 4: Neck Dissection Material Scores

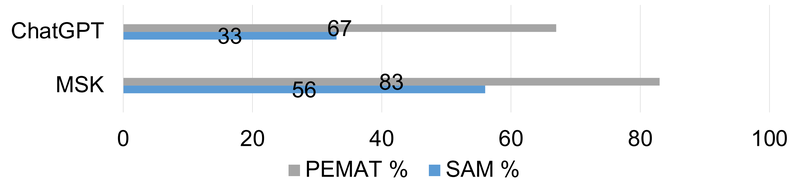
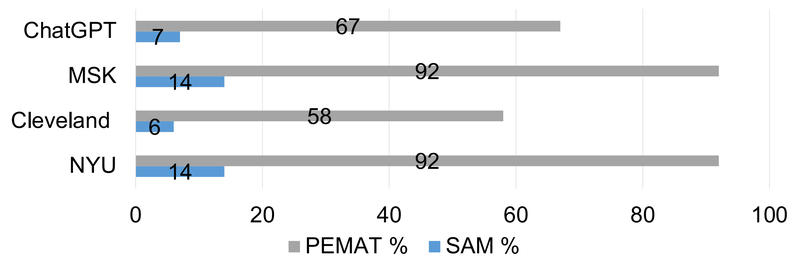


Figure 5: Glossectomy Material Scores



RESULTS

- 1 of 4 sites had comprehensive online material for neck dissection and 3 of 4 had material for glossectomy. Comparison data for the other 3 clinical scenarios are shown in Figures 1-5.
- ChatGPT performed at or above the median score for all scenarios.
- ChatGPT responses all came back as college level readability (5/5), similar only to 1 other institution's materials.
- ChatGPT received the lowest scores possible when material interaction was assessed on the SAM (0/2 all 5 times) and when graphics were assessed on both the SAM and PEMAT.
- ChatGPT performed similar to other materials in text organization, content, and writing style.

DISCUSSION

- ChatGPT's performance in counseling and educating patients about their surgery is overall comparable to current online material, making it an acceptable resource for at-home use, as well as for providers wishing to develop educational material.
- ChatGPT's stagnant responses did not perform as well in the interaction category for the SAM. However, the ability to interact with the user when prompted is unique and not adequately assessed on the PEMAT or SAM. This is a valuable property of ChatGPT and should be considered an asset in its ability to provide patient education.
- With further refinement of AI, we anticipate basic perioperative counseling can be performed accurately and effectively. Refinements, such as those seen visual-generating AI, can be used to enhance the patient education experience. AI can be used to patient concerns during the perioperative period, and future iterations can add to this capability.

REFERENCES

- Liu, P. et al. Application of Artificial Intelligence in Medicine: An Overview. Curr. Med. Sci. 41, 1105–1115 (2021).
- Vanian, J. Why tech insiders are so excited about ChatGPT, a chatbot that answers questions and writes essays. CNBC https://www.cnbc.com/2022/12/13/chatgpt-is-a-new-ai-chatbot-that-can-answer-questions-and-write-essays.html (2022)

CONTACT:
Rusha Patel, MD, FACS
800 Stanton L. Young Blvd, Email: rusha-patel@ouhsc.edu, Ph: 405-271-4000